SIGNAL INNOVATIONS GROUP

Quarterly Report

2/25/2006-05/24/2006

This research effort is focusing on development of adaptive sensing algorithms for asymmetric threats. The algorithms are particularly targeted towards use with data from multi-camera video surveillance systems. We are not attempting to model the infinite class of asymmetric targets, since these are generally unpredictable and therefore limited if any *a priori* sensor data are available for these threats. Rather, we model normal or typical behavior using statistical algorithms.

Key challenges that are being addressed in the current effort:

- Statistical characterization of background events
- Tracking of foreground objects
- Statistical characterization of object dynamics (via HMMs), in the presence of occlusions and uncertainty of object pose and motion
- Multi-aspect, multi-camera target recognition

A principal requirement for anomalous event detection in video data is to separate foreground object activity from the background scene. SIG has previously investigated using an inter-frame difference approach that yields high intensity pixel values in the vicinity of dynamic object motion. While the inter-frame difference is computationally efficient, it is ineffective at highlighting objects that are temporarily at rest and is highly sensitive to natural background motion not related to activity of interest such as tree and leaf motion. SIG is currently employing a statistical background model using Gaussian mixture (GMM), with the background image corresponding to a sum of Gaussian random variables that represent the statistical variations of the background pixels. The GMM estimates parameters for each pixel in RGB space yielding a likelihood that the pixel belongs to either the background or a set of foreground objects. These parameters are updated using a highly efficient real time implementation of the expectation maximization (EM) algorithm. In order to accommodate dynamic background modeling, statistics of the scene that vary over time are integrated into a unique model update sub-system to refine the parameter estimation.

4309 Emperor Blvd., SUITE 350, DURHAM, NC 27703
FAX: 919.287.2398 WEB: WWW.SIGINNOVATIONS.COM

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **24 MAY 2006**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**
4. TITLE AND SUBTITLE **Quarterly Report 2/25/2006-05/24/2006**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Signal Innovations Group 4309 Emperor Blvd., Suite 350 Durham, NC 27703**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

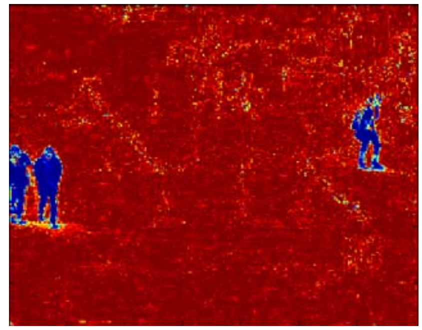
14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**SAR**	**5**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

a) Original video image with multiple static occlusions

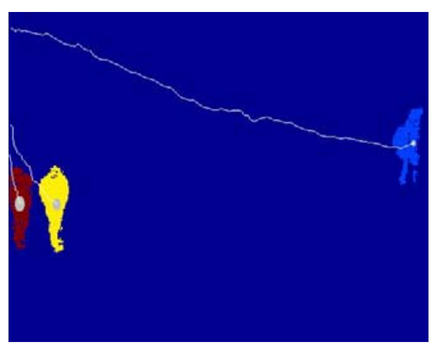


b) likelihood of background association; contiguous blue regions are foreground objects

Figure 1 Background likelihood created by dynamic background sub-system

SIG has also investigated nonlinear object ID and tracking methods. The objects within a scene are characterized via a feature-based representation of each object. Kalman filtering and particles filters have been implemented to track object position and velocity through the video sequence. A point of reference for each object (i.e. center of mass) is tracked through video sequence. Given an adequate frame rate, greater than 3 frames per second, we can assume that this motion is linear. Kalman filters provide a closed-form solution to track position and velocity given Gaussian noise and produces likelihood values of the given objects in the scene. The values are then sent to the SMA for further processing and passed back via a feedback loop to the update sub-system to further enhance feature extraction. Thus temporal information from past frames can be exploited to mitigate the effects of abrupt lighting changes and occlusions. This methodology combines the best aspects of both GMM and MRF into a single compact, analytical algorithm.

a) Original image



b) Object association and dynamic tracking

Figure 2 Object association and dynamic tracking

Occlusions are handled via a statistical shape model, which adaptively learns likely spatio-temporal associations with pixels associated with an object with respect to an object reference point (e.g. centroid). When an occlusion occurs, a pixel in the vicinity of the occlusion is capable of being associated with multiple objects as illustrated below.

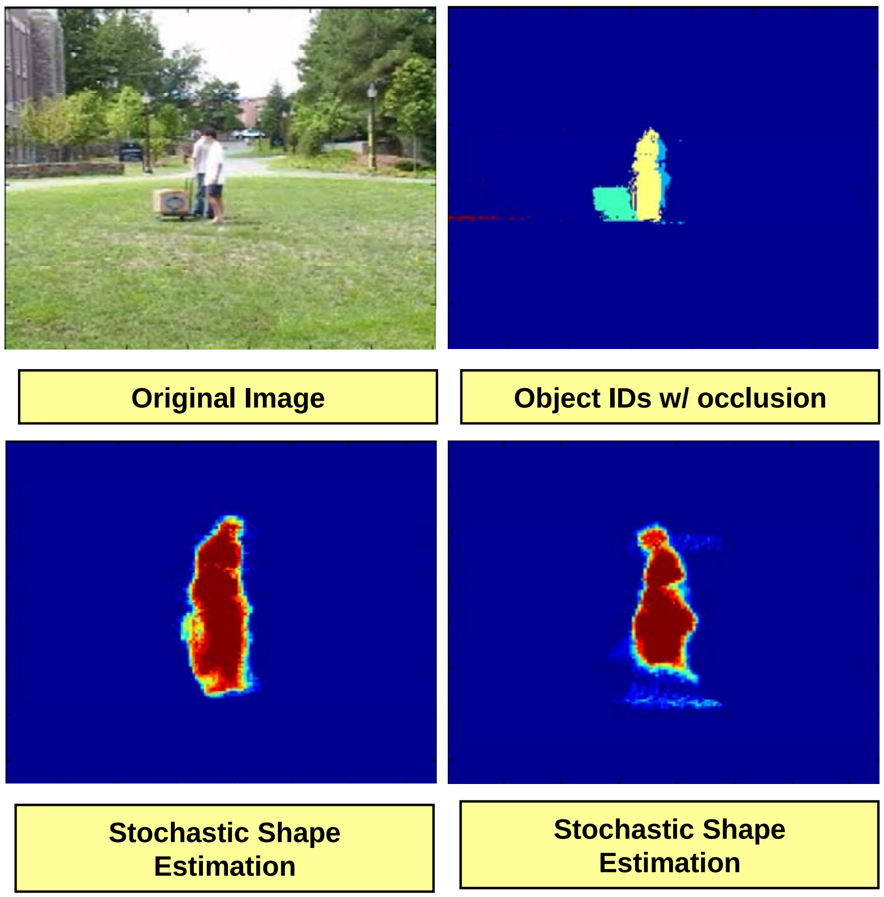


Figure 3 Shape estimation and occlusion mitigation

The sequential data characteristic of targets in video are modeled using hidden Markov models (HMMs). There are two key challenges that must be addressed when performing such analyses: (i) the different classes of typical behavior, such as individuals, groups and vehicles will require different types of HMMs, and therefore the number of HMMs is not known *a priori*; and (ii) the characteristics of each of the behavior-dependent HMMs are not known *a priori*, for example the number of states required in the respective HMMs.

To address these challenges we are employing new techniques in statistical analysis, termed Dirichlet processes (DPs). The HMM may be viewed as a statistical density function on sequential data. Our objective is to learn the number of different types of HMMs (different classes of normal/typical behavior) as well as the properties of each.

The DP construction constitutes a probability density function on the HMMs; in other words, DP is a probability density function on the particular HMMs that are appropriate for characterizing a given time-evolving scene. Using the DP setting, we have developed a framework that autonomously indexes different forms of normal video data, while simultaneously learning the associated HMM representation of each class of data. When performing surveillance, any given sequence of data from the video is then submitted to the DP HMM mixture model, and if the activity appears to be anomalous it yields a low likelihood of being characteristic of typical/normal behavior. Data that appear to be atypical are sent to an analyst for evaluation. If the analyst characterizes the data as being a non-threat, then the HMM representation is updated with the introduction of a new class of typical behavior. In this manner the algorithm and video system learns over time what behavior is deemed to be typical, detecting those activities that are unusual.

To date, SIG has collected a large quantity of video data with which to perform these analyses and algorithm development. In addition, fixed cameras are now being deployed looking out of the SIG facility, for further video analysis and algorithm refinement. We are on target to deploy a system for testing at China Lake, during the second year of this program. The fundamental system-level concepts are being developed in Matlab and tested on video data collected by SIG, or provided to SIG by third parties.

The SIG research is being performed in collaboration with Lockheed Martin (LM). Earlier LM work performed under NAVAIR/ONR funding had focused on recognizing targets using views from multiple sensors/look directions, given that they had been detected. The current effort is seeking to improve the target detection process using multiple sensor views. During this period, Lockheed Martin has formulated a methodology for detecting objects using multiple sources of video imagery, and a basic framework for the approach has been developed. The basic concept is shown in Figure 1.

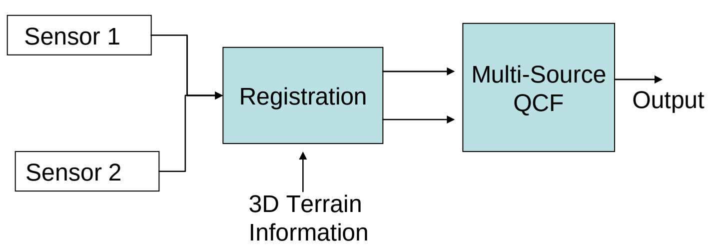


Figure 4: Multi-aspect QCF ATR architecture

As shown in the figure, assume that the scene is observed simultaneously by two video cameras at different locations. The knowledge of the sensor parameters and look direction are used to morph one view to match the other. Although affine transforms may suffice for relatively flat locations, a 3D terrain data base (such as the SRTM data) may be used to achieve registration in the more general case.

Assuming that registration between the views has been achieved, the separation in the detection metric φ for Target and Clutter is

$$E_T\{\varphi\} - E_C\{\varphi\} = \begin{bmatrix}\mathbf{x}^T & \mathbf{y}^T\end{bmatrix}\begin{bmatrix}\mathbf{T}_{11} - \mathbf{C}_{11} & \mathbf{T}_{12} - \mathbf{C}_{12} \\ \mathbf{T}_{21} - \mathbf{C}_{21} & \mathbf{T}_{22} - \mathbf{C}_{22}\end{bmatrix}\begin{bmatrix}\mathbf{x} \\ \mathbf{y}\end{bmatrix} = \begin{bmatrix}\mathbf{x}^T & \mathbf{y}^T\end{bmatrix}\mathbf{R}\begin{bmatrix}\mathbf{x} \\ \mathbf{y}\end{bmatrix}$$

where **x** and **y** are the simultaneous registered images obtained from two sensors, $\mathbf{T}_{ij}$ is the correlation of the target signatures between the sensor *i* and sensor *j*, and $\mathbf{C}_{ij}$ is the same for clutter. The set of eigenvectors of the matrix **R** are the 2-channel QCF kernels depicted in Figure 2.

The required matrices will be estimated from training data of both targets and clutter. Since registration is not expected to be perfect, some of the imperfections will be modeled in the training process by purposefully introducing parallax effects, and spatial offsets between training images of the two sensors. Our next effort is to collect and ground truth data using two stationary cameras, and running simulations to assess the performance gain, if any, of the proposed approach relative to a single sensor system.

In this reporting period, the quadratic correlation filters have been extended to a generalized polynomial correlation filter. The technique has been developed and initially assessed on imagery containing tactical vehicles and civilian vehicles. In particular, we desire that the algorithm be robust to, and in fact exploit, multi-aspect interrogation of the target vehicle. Results on initial data collections are presented in the **Appendix A** Attachment. Such methods will dovetail into SIG's attempts to model the identity and dynamics of vehicles when viewed from multiple aspects and from multiple cameras.